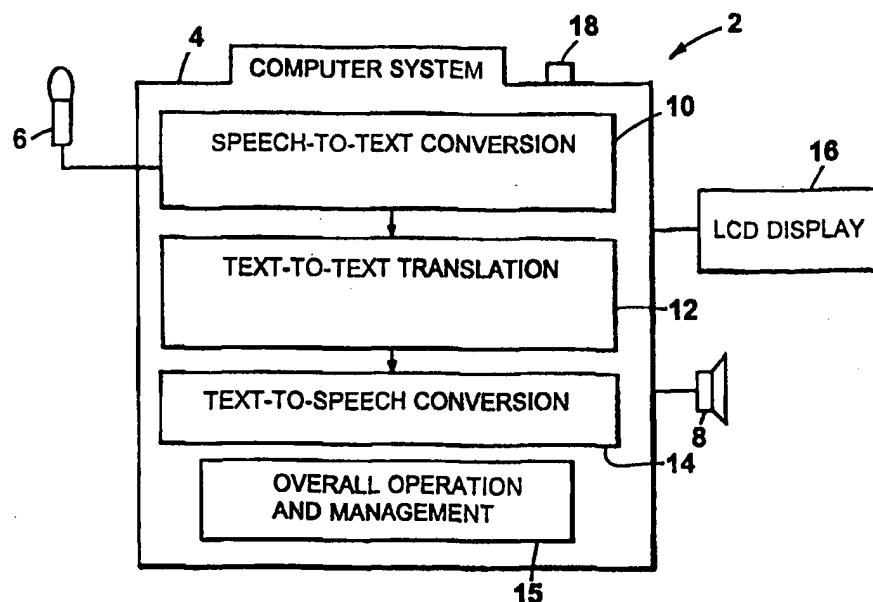




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G10L 3/00	A1	(11) International Publication Number: WO 99/46762 (43) International Publication Date: 16 September 1999 (16.09.99)
<p>(21) International Application Number: PCT/US99/05058</p> <p>(22) International Filing Date: 9 March 1999 (09.03.99)</p> <p>(30) Priority Data: 60/077,263 9 March 1998 (09.03.98) US 09/097,216 12 June 1998 (12.06.98) US</p> <p>(71) Applicant: KELVIN LP [US/US]; 10 Hub Drive, Melville, NY 11747 (US).</p> <p>(72) Inventor: HADAR, Avinoam; 3 Magnolia Lane, Woodbury, NY 11797 (US).</p> <p>(74) Agent: BORODACH, Samuel; Fish & Richardson P.C., Suite 2800, 45 Rockefeller Plaza, New York, NY 10111 (US).</p>		<p>(81) Designated States: CA, IL, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p>Published <i>With international search report.</i></p>

(54) Title: AUTOMATIC SPEECH TRANSLATOR



(57) Abstract

A speech translator (2) includes a microphone (6) and a speech-to-text processor (10) configured to receive speech signals in a first language and to convert the speech signals to a first text file in the first language. The first text file is converted to a second text file in a second language (12). The speech translator includes a speech synthesizer (14) which converts the second text file to speech signals in the second language. An amplifier receives and amplifies the speech signals in the second language.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

- 1 -

AUTOMATIC SPEECH TRANSLATORBackground

5 The invention relates generally to automatic speech translation.

 In recent years, travel between countries for the purposes of business and pleasure has increased. Unfortunately, visitors to a foreign country often do not
10 speak or understand the native language(s), and, therefore, conversing with inhabitants of the country which they are visiting can be difficult. The difficulties associated with reading documents in a foreign language, and conversing with another person who
15 does not understand the speaker's language and whose language the speaker does not understand, create a language barrier. Such a language barrier impedes business as well as pleasure while visiting or travelling in a foreign country.

Summary

20 In general, according to one aspect, a speech translator includes a microphone and a processor configured to receive speech signals sensed by the microphone where the speech signals represent speech in a
25 first language. The processor is further configured to convert the speech signals to a first text file in the first language and to convert the first text file to a second text file in a second language. The translator

- 2 -

includes a speech synthesizer which converts the second text file to speech signals in the second language. An amplifier receives and amplifies the speech signals in the second language.

5 The translator can be a stand-alone portable unit or can be implemented using, for example, a personal computer. The translator can translate speech between two or more languages so that persons speaking different languages can converse without the need for a human
10 translator. The translator also can be used, for example, for other purposes, such as phonic and pronunciation training in a single language or in different languages.

 The invention also includes a method of performing
15 speech translation. Exemplary implementations of the translator and method are discussed in greater detail below. Various implementations include one or more of the following advantages. The translator can be compact, making it particularly convenient for use during
20 travel abroad, for example, at business meetings. In some implementations, the translator can be incorporated as part of a headphone set to facilitate its use in theaters or in a lecture hall. In addition, the translator can be versatile, allowing the user to select
25 from a menu the languages in which the parties will converse. Some implementations permit bi-directional translation.

 Other features and advantages will be apparent from the detailed description, accompanying drawings and
30 the claims.

Brief Description of the Drawings

FIG. 1 is a block diagram illustrating one implementation of a speech translator according to the invention.

- 3 -

FIG. 2 illustrates a portable speech translator according to the invention.

- 4 -

Detailed Description

As shown in FIG. 1, an automatic speech translator 2 is designed to allow persons speaking different languages, such as English and French, or English and Japanese, to communicate with one another even if one person, for example, does not speak or understand the other person's language. The translator 2 can be implemented, for example, using a laptop computer.

The speech translator 2 includes a computer or other processor 4, at least one microphone 6, an amplifier such as a speaker 8, and a monitor 16. In various implementations, the computer 4 can be a personal computer configured with the appropriate software, or a microprocessor programmed with dedicated software. For example, a Pentium 300 device is suitable for some implementations. While the microphone 6 and the speaker 8 are illustrated as discrete components of the translator 2, in some implementations, the microphone and the speaker can be integrated with the translator as part of a single integrated unit. An operating and management software module 15 controls the overall operation of the computer 4 and the interaction between the various components of the translator 2.

The translator 2 is designed to translate speech from a first language, such as English, to a second language, such as French. In some implementations, the first and second languages are selected, for example, at the time of manufacture, and the translator 2 is configured to convert speech only from the first language to the second language. In other implementations, however, the translator 2 can include a switch or menu which allows the user to select either the first language, the second language, or both, from among two or more options.

- 5 -

The computer 4 includes a speech-to-text conversion unit 10, a text-to-text translation unit 12, and a text-to-speech conversion unit 14. The units 10, 12, 14 can be implemented in hardware, software, or a combination of both hardware and software. In general, speech signals received by the microphone 6 are directed to the speech-to-text unit 10, for example, through a sound card associated with the computer 4. The unit 10 converts the received signals into a first digital text file in the same language as the received speech. In one particular implementation, the speech-to-text unit 10 includes a continuous speech recognition software module, such as the Naturally Speaking™ system available from Dragon Systems, Inc. The text can be displayed on the display 16 to permit the person speaking to confirm that the speech-to-text translation is accurate. Alternatively, the translator 2 can re-convert the text output from the speech-to-text unit 10 to audio signals to allow the speaker to confirm the accuracy of the translation from speech to text.

The computer 4 further is configured to convert the text from the first language to the second language using the text-to-text translation unit 12. In one particular implementation, the translation unit 12 can include a software module such as the Power Translator™ available from Global Link. In general, proper names, including names of persons or places, should be transliterated rather than translated. The text-to-text translation unit 12 forms a second text file which represents the received speech in the second language. In some implementations, the translated text in the second-language is displayed on the display.

The second text file then is converted to speech in the second language by the text-to-speech conversion unit 14. The text-to-speech unit 14 includes a speech

- 6 -

synthesizer, and, in one particular implementation, can include a software module such as Mac in Talk™ available from Macintosh or TrueVoice™ available from LearnOut & Hauspie. Other suitable text-to-speech software modules
5 are available, for example, from International Business Machine, Inc. The signals generated by the text-to-speech conversion unit 14 are sent to the speaker 8 which amplifies the received signals so that they can be heard by persons in the vicinity of the translator 2.

10 To illustrate the operation of the translator 2, it is assumed that the translator 2 is configured to translate speech from English to French. In some cases, prior to using the translator 2, it may be desirable or necessary for a person to speak sample words or phrases
15 to train the translator and to allow it to recognize a particular person's accent.

Once the training process (if any) has been completed, a first person would speak in English into the microphone 6. The person might say, for example, "Where
20 is the bus?" The speech-to-text unit 10 would receive digital signals representing the speech and would form a first text file corresponding to that speech. In addition, the text "Where is the bus?" would appear on the display 16. The text-to-text translation unit 12
25 would convert the first text file to a second text file in French, and the text "Où est l'autobus?" would appear on the display 16. The text-to-speech unit 14 would generate sounds corresponding to the sentence "Où est l'autobus?" which would be amplified by the speaker 8 so
30 that one or more other persons in the vicinity of the translator 2 would hear the translated sentence in French. In some implementations, the translator 2 is programmed to perform the translation word by word, whereas in other implementations, the translation is
35 performed, for example, sentence by sentence.

- 7 -

In some implementations, the translator 2 can be used for converting speech from the second language (e.g. French) to the first language (e.g., English) as well. In such an implementation, the various units 10, 12, 14 are configured to handle the conversions and translations from the first language to the second language as well as from the second language to the first language. The translator 2 is programmed to be in one of two modes. In the first mode, the translator 2 assumes that the received speech signals correspond to speech in the first language for translation to the second language, whereas in the second mode, the translator 2 assumes that the received speech signals correspond to speech in the second language for translation to the first language. A manual switch or button 18 can be provided on the exterior of the translator 2 to switch between the first and second modes. Alternatively, a two-microphone sound card can be provided which allows the translator automatically to identify each language and perform the translation without the need to press the switch 18.

FIG. 2 illustrates additional features of an automatic speech translator 20. In general, the translator 20 is configured and programmed to translate speech between selected languages in a manner similar to the translator 2 described above. The translator 20 is a portable unit which can be held, for example, in a person's hand. The translator 20 includes a power switch 22 for turning the unit on and off, a volume switch or knob 24 for controlling the volume of generated speech. The translator 20 also includes a menu display 38 from one which one of several options, including at least either the first or second language, can be selected by using a selection button or knob 26. The selection button or knob 26 allows one to scroll through the available options and to select one of the options from

- 8 -

the menu. For example, in one implementation, the translator 20 is capable of translating speech between any two of the following languages: English, Japanese, French, German, Italian, Russian or Chinese. Of course, 5 in other implementations, additional or different languages are available. The menu display 38 can be a liquid crystal display (LCD).

The translator 20 includes multiple input ports 28A, 28B for connection to respective microphones. For 10 example, one person can use a microphone connected to the input port 28A. A second person can use a microphone connected to the input port 28B. The translator 20 also has a speaker 36 through which the translated speech can be heard and a liquid crystal display (LCD) 34 on which 15 the text of the original and translated speech can be displayed. The display 34 can be folded onto the top of the translator 20 when the unit is not being used to provide additional compactness. As in the implementation of FIG. 1, a switch or button 32 can be provided on the 20 exterior of the translator 20 to switch between different modes which perform translations respectively from a first language to a second language and vice-versa.

The translator 20 includes multiple output ports 30A, 30B for connection to respective headphone sets. 25 For example, the person using a microphone connected to the input port 28A, can use a headphone set connected to the output port 30A. Similarly, the person using a microphone connected to the input port 28B can use a headphone set connected to the output port 30B. When 30 headphone sets are connected to the output ports 30A, 30B, the translated speech signals are heard only by the persons wearing the headphone sets. This feature can provide additional privacy to the parties carrying on the conversation and can lessen disturbances to the parties 35 from other noise in the vicinity. Optionally, the

- 9 -

translator can include a port 40 for connection to other input/output devices, such as a keyboard, hand-writing recognition devices, or a printer. Thus, for example, in some implementations, the translator 2 saves the entire
5 dialogue which subsequently can be printed out.

The translator 20 also can include holders for the microphone and headphone sets when they are not in use. The holders can be attached to the side of the translator 20. Additionally, the translator 20 can include a
10 shoulder strap 42 for carrying the translator.

While the foregoing implementations have been described in the context of translating speech from one language to a different second language, the translators 2 and 20 also can be used phonic and pronunciation
15 training. Thus, if a person wishes to improve his pronunciation, for example, of a foreign language, the translator can be configured so that the selected first and second languages are the same.

In various implementations, the translator 2 (or
20 20) can be used to translate speech from television, radio, telephone or other speech sources as well as directly from another person. The translator 2 (or 20) also can be connected to a telephone line to permit conversations to take place, for example, across the
25 Internet. Additionally, a compact disc (CD) drive can be included as part of the translator.

Other implementations are within the scope of the following claims.

- 10 -

What is claimed is:

1. A speech translator comprising:
 - a microphone;
 - a processor configured to receive speech signals
 - 5 sensed by the microphone, wherein the speech signals represent speech in a first language, and wherein the processor is further configured to convert the speech signals to a first text file in the first language, to convert the first text file to a second text file in a
 - 10 second language;
 - a speech synthesizer for converting the second text file to speech signals in the second language; and
 - an amplifier for receiving and amplifying the speech signals in the second language.
- 15 2. The speech translator of claim 1 wherein the first and second languages are different.
3. The speech translator of claim 2 further including:
 - a display for displaying a menu of languages from
 - 20 which at least one of the first and second languages can be selected; and
 - means for selecting an option displayed on the menu.
4. The speech translator of claim 3 wherein the
- 25 display is a liquid crystal display.
5. The speech translator of claim 2 further including a display for displaying text corresponding to the first or second text files.

- 11 -

6. The speech translator of claim 5 wherein the display can be folded over a top of the translator when the translator is not in use.

7. The speech translator of claim 2 having:

5 a first mode in which speech signals in the first language are converted into speech signals in the second language;

a second mode in which speech signals in the second language are converted into speech signals in the
10 first language;

wherein the speech translator further includes means for switching between the first and second modes.

8. The speech translator of claim 7 wherein the means for switching includes a two-microphone sound card.

15 9. The speech translator of claim 7 wherein the means for switching includes a manual switch.

10. The speech translator of claim 2 further including:

output ports for connection to headphone sets
20 wherein the output ports are coupled to the amplifier.

11. A method of performing speech translation, the method comprising:

receiving signals representing speech in a first language;

25 converting the received signals to a first text file in the first language;

translating the first text file to a second text file in a second language;

generating speech signals corresponding to the
30 second text file.

- 12 -

12. The method of claim 11 wherein the first and second languages are different..

13. The method of claim 12 further including:
amplifying the speech signals corresponding to the
5 second text file.

14. The method of claim 12 further including:
displaying a menu of languages from which at least
one of the first and second languages can be selected.

15. The method of claim 12 further including:
10 displaying the text corresponding to the first or
second text files.

16. The method of claim 12 further including:
switching between first and second modes, wherein
in the first mode speech signals in the first language
15 are converted into speech signals in the second language,
and wherein in the second mode speech signals in the
second language are converted into speech signals in the
first language.

